PRESCRIPTIVE ANALYTICS BASED DATA WAREHOUSE CLUSTER NODE OPTIMIZATION STACK FOR CLOUD COMPUTING

Technical Field

[001] This disclosure relates to storage cluster node optimization via a prescriptive analytics based data warehouse cluster node optimization stack.

Background

[002] Rapid advances in communications and storage technologies, driven by immense customer demand, have resulted in widespread adoption of cloud systems for managing large data payloads, distributed computing, and record systems. As one example, modern enterprise systems, for example those relying on object storage architectures, presently maintain data records many petabytes in size in the cloud. Improvements in tools for cloud resource allocation, utilization, and consumption prediction will further enhance the capabilities of cloud computing systems.

Summary of the Invention

[002a] According to a first aspect of the present invention, there is provided a system comprising: network interface circuitry configured to: receive utilization tracking data for a selected computing resource cluster of computing resource nodes; receive network traffic data for the selected computing resource cluster; and send a cluster node optimization (CNO) token to a host interface, the host interface configured to control cluster node optimization for at least the selected computing resource cluster; optimization circuitry in data communication with the network interface circuitry, the optimization circuitry configured to execute a cluster node optimization (CNO) stack, the CNO stack executable to: at an input layer of the CNO stack: obtain, via the network interface circuitry, the utilization tracking data and the network traffic data; and store the utilization tracking data and the network traffic data via a storage operation at a data-staging layer of the CNO stack; at a configuration layer of the CNO stack: determine a CNO interval; and determine optimization metric thresholds; at a prescriptive engine layer of the CNO stack: select target computing resource nodes from the selected computing resource cluster based on the utilization tracking data,

the optimization metric thresholds, and the CNO interval; execute a prediction engine to predict respective storage utilizations over a next operation cycle for the target computing resource nodes; generate an aggregated storage utilization prediction for the selected computing resource cluster based on the predicted storage utilizations; determine a network traffic coefficient for the selected computing resource cluster based on the network traffic data; perform a cluster determination to execute a cluster node optimization for the selected computing resource cluster based on the aggregated storage utilization prediction and the network traffic coefficient; and generate the CNO token based on the cluster determination.

[002b] According to a second aspect of the present invention, there is provided a method comprising: at network interface circuitry: receiving utilization tracking data for a selected computing resource cluster of computing resource nodes; and receiving network traffic data for the selected computing resource cluster; at optimization circuitry in data communication with the network interface circuitry, the optimization circuitry executing a cluster node optimization (CNO) stack by: obtaining, at an input layer of the CNO stack, the utilization tracking data and the network traffic data via the network interface circuitry; storing, at the input layer, the utilization tracking data and the network traffic data via a storage operation at a data-staging layer of the CNO stack; determining, at a configuration layer of the CNO stack, a CNO interval; determining, at the configuration layer, optimization metric thresholds; selecting, at a prescriptive engine layer of the CNO stack, target computing resource nodes from the selected computing resource cluster based on the utilization tracking data, the optimization metric thresholds, and the CNO interval; predicting, at the prescriptive engine layer, with a prediction engine, respective storage utilizations over a next operation cycle for the target computing resource nodes; generating, at the prescriptive engine layer, an aggregated storage utilization prediction for the selected computing resource cluster based on the predicted storage utilizations; calculating, at the prescriptive engine layer, a network traffic coefficient for the selected computing resource cluster based on the network traffic data; determining, at the prescriptive engine layer, to execute a cluster node optimization for the selected computing resource cluster based on the aggregated storage utilization prediction and the network traffic coefficient; and generating, at the prescriptive engine layer, a CNO token based on the determination to execute the cluster node optimization; and

sending, via the network interface circuitry, the CNO token to a host interface configured to control cluster node optimization for at least the selected computing resource cluster.

[002c] According to a third aspect of the present invention, there is provided a product including: machine-readable media other than a transitory signal; and instructions stored on the machine-readable media, the instructions configured to, when executed, cause a machine to: at network interface circuitry: receive utilization tracking data for a selected computing resource cluster of computing resource nodes; and receive network traffic data for the selected computing resource cluster; at optimization circuitry in data communication with the network interface circuitry, the optimization circuitry executing a cluster node optimization (CNO) stack: obtain, at an input layer of the CNO stack, the utilization tracking data and the network traffic data via the network interface circuitry; store, at the input layer, the utilization tracking data and the network traffic data via a storage operation at a data-staging layer of the CNO stack; determine, at a configuration layer of the CNO stack, a CNO interval; determine, at the configuration layer, optimization metric thresholds; select, at a prescriptive engine layer of the CNO stack, target computing resource nodes from the selected computing resource cluster based on the utilization tracking data, the optimization metric thresholds, and the CNO interval; predict, at the prescriptive engine layer, with a prediction engine, respective storage utilizations over a next operation cycle for the target computing resource nodes; generate, at the prescriptive engine layer, an aggregated storage utilization prediction for the selected computing resource cluster based on the predicted storage utilizations; calculate, at the prescriptive engine layer, a network traffic coefficient for the selected computing resource cluster based on the network traffic data; determine, at the prescriptive engine layer, to execute a cluster node optimization for the selected computing resource cluster based on the aggregated storage utilization prediction and the network traffic coefficient; and generate, at the prescriptive engine layer, a CNO token based on the determination to execute the cluster node optimization; and send, with network interface circuitry, the CNO token to a host interface configured to control cluster node optimization for at

least the selected computing resource cluster.

Brief Description of the Drawings

[003] The present disclosure may be better understood with reference to the following drawings and description. The components in the figures are not necessarily to scale. Moreover, in the figures, like-referenced numerals designate corresponding parts throughout the different views.

[004] Figure 1 shows an example multiple-layer cluster node optimization stack.

[005] Figure 2 shows example cluster node optimization logic.

[006] Figure 3 shows an example specific execution implementation for a prescriptive engine layer of the cluster node optimization stack in accordance with an embodiment.

[007] Figure 4 shows an example cluster node optimization control interface.

[008] Figure 5 shows a second example cluster node optimization control interface.

[009] Figure 6 shows an example interaction logic with a cluster node optimization control interface in accordance with an embodiment.